Stochastic search on decision trees

Ilona Krasznahorkay

The construction of decision trees is an often used and easily applied way of supervised learning. The goal is the prediction of a binary target variable depending on many predictor variables. The algorithm divides the field of predictors along a predictor variable one after another. The aim is to find a partition where the target variables are the most homogeneous. I modified the deterministic CART algorithm developed by Breiman and others [1]– which aims for the minimizing of a concave risk function defined on the partitions generated by the trees. I used the Markov Chain Monte Carlo method so doing stochastic searches on the set of decision trees. (It was first proposed in a Bayesian framework by Chipman and others [2].) By empirical experience finding the optimal tree with this technique is much more effective than the former deterministic methods.

In my talk the main result is the examination of MCMC type stochastic searches on decision trees. I prove the geometric ergodicity of the constucted Markov chains using the drift-criterion technique. I prove exact results on the mixing time of the above Markov chains using different methods like canonical paths, conductance and coupling, see Jerrum, [4] and Jerrum & Sinclair, [5].

I also examine the cost-complexity risk function, where the penalty term depends on the number of the leaves of the tree. In this case the associated Gibbs distribution is studied.

References

[1] BREIMAN, L., FRIEDMAN, J.H., OLSEN, A.O., STONE, C.J.: *Classification and Regression Trees.* Wadsworth International Group, 1984.

[2] CHIPMAN, H.A., GEORGE, E.I., MCCULLOCH, R.E.: *Bayesian CART Model Search.* JASA, Vol. 93. Num. 443. (1998), 935-960.

[3] MEYN, S.P., TWEEDIE, R.L.: *Markov Chains and Stochastic Stability.* Springer, 1993.

[4] JERRUM, M., *Mathematical foundations of the Markov chain Monte Carlo method.* Probabilistic methods for algoritmic discrete mathematics, 116-165, Algoriths Combin., 16, Springer, Berlin, 1998.

[5] JERRUM, M., SINCLAIR, A., *The Markov chain Monte Carlo method: An approach to approximate counting and integration.* In: Approximation Algorithm for NP-hard Problems, (Dorit Hochbaum, ed.) PWS, 1996.